

다양한 환경에 강건한 RGB 영상 기반 보행 분석

안지민 · 정겨운 · 신동인 · 원건 · 박종범

전자부품연구원

Robust RGB image-based gait analysis in various environment

Ahn Ji-min · Jeung Gyeo-wun · Shin Dong-in · Won Geon · Park Jong-beom

Korea Electronics Technology Institute

E-mail : hihijimin@keti.re.kr, jbpark@keti.re.kr*

요 약

본 논문은 RGB 영상 이용하여 하지 움직임에 대한 분석을 다룬다. 딥러닝 접근방법인 객체 인식 Segmentation 알고리즘과 자세 검출 알고리즘을 융합한 방법과 BMC(Background Model Challenge)을 활용하여 RGB 영상을 보행 분석 요소로 사용하였다. 본 연구에서 제시한 영상 보행 분석은 보행패턴 인식과 비정상적인 보행 등의 분류를 위한 변수로서 활용할 수 있을 것으로 판단된다.

ABSTRACT

This paper deals with the analysis of leg motion using RGB image. We used RGB image as gait analysis element by using BMC(Background Model Challenge) method and by using combining object recognition segmentation algorithm and attitude detection algorithm. It is considered that gait analysis incorporating image can be used as a parameter for classification of gait pattern recognition and abnormal gait.

키워드

gait parameter, skeleton, deep learning

I. 서 론

인간의 보행은 어린아이부터 노인까지 매우 다양한 형태로 나타난다. 노인성 편마비와 같은 특정한 질환이나 스포츠 선수들의 부상 등 보행의 경우 보행 불균형으로 인해 어려움을 겪고 있다. 보행 불균형에서 나타나는 큰 특징은 보행 시 다리의 짧은 지지 시간과 작은 지상반력(Ground reaction force)이 있다[1].

Griffin[2]은 반신 마비 환자의 걸음을 관찰하여 불균형 특징을 확인 하였다. 이러한 운동학적 요소 방법은 사람의 골격특성을 고려하지 않은 단점이 있다.

본 논문에서는 보행 시 사람의 골격 형태를 파악하기 위한 방법으로 두 가지 접근법을 비교한다. 첫 번째는 BMC(Background Model Challenge)을 이용한 skeleton 추출법[3]이며 두 번째는 객체를 인식하여 segmentation 해주는 Tusimple DUC 알고리즘과 사람 검출 자세 측정 라이브러리인 Openpose를 융합하여 적용한 접근법이다.

이 두 접근법의 비교를 통해 다양한 환경과 보

행 시 불균형 정도를 파악하기 위한 파라미터로서 Wang[4]가 제안한 Tusimple DUC 와 CaoZ[5]가 제안한 Openpose 융합한 방법이 보다 적합한 접근 방법임을 제안한다.

II. 본 론

2.1. 보행 측정 장치구성 및 측정 방식

Microsoft 사의 Kinect v2 센서를 사용하여 보행 데이터를 취득하였다.

보행 대상자와 카메라 간 거리는 210cm로 설치하였으며, 객체의 총 보행 거리는 230cm(일반 건강 한 일반인 기준 약 7~8 걸음) 이다.

2.2. 보행 파라미터 분석

2.2.1. BMC(Background Model Challenge)을 이용한 방법

전처리 과정으로서 7x7 픽셀단위로 메디안 필터를 사용하여 입력영상의 잡음 제거 시켰다. 그런 후, 움직이는 객체를 탐지하고 추적하기 위하여 두

영상의 차를 이용한 배경제거 알고리즘인 Background Subtraction 을 적용하여 실루엣 영상을 추출하였다. 실루엣 영상에서 각 픽셀의 값 범위는 0~255이며 각 픽셀에서 0과 255 외의 값을 제거 하여 2진 영상으로 만든 다음 객체의 경계선을 뽑아내었다. 그런 후, 이웃한 픽셀들 중 최소 픽셀값을 현재 픽셀값으로 뽑아내는 과정(침식)과 반대로 이웃한 화소들 중 최대 픽셀값을 현재 픽셀값으로 대체하는 과정(팽창)을 거쳐 이 두 픽셀 값의 차를 계산하여 보행하는 객체의 skeleton을 추출 하였다 [3]. 그림 1은 RGB 입력 영상에 대한 영상처리 단계별 적용한 결과이다.

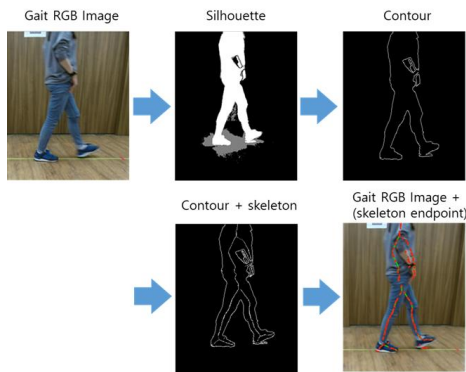


그림 1. BMC 이용한 보행 자세 골격 구조 추출 결과 (Gait RGB Image - Silhouette - Contour - Contour+Skeleton - Gait RGB Image + Skeleton endpoint)

2.2.3 객체 인식 Segmentation 알고리즘과 자세 검출 알고리즘을 융합한 방법

영상에서 보행인 객체만을 추출하기 위해 Tusimple DUC 알고리즘을 적용하였다[4]. 이는 DeepLab-V2 ResNet-101 네트워크를 기본모델로 사용한 것으로 보행 데이터를 입력으로 두고 출력층에는 분류 할 객체를 카테고리 별(사람, 건물, 배경 등)로 분류할 수 있도록 지정하였고 교차 엔트로피(Cross-Entropy) 형태의 오차 함수를 사용하여 출력층에서 활성화 함수의 도함수에 의한 영향을 제거 하도록 하였다. 출력 맵에서 모든 픽셀 위치에 결과가 합계되며 SGD (Standard Stochastic Gradient Descent)을 사용하여 결과를 최적화 시켰다[4]. 그림 2의 가운데 이미지는 본 과정의 결과이다.

위 과정에서 입력영상의 객체 segmentation을 추출한 후, 2분기 다중 스테이지 CNN (Convolutional neural network) 모델을 통과시켜 관절이 있을 것 같은 위치를 표현하는 Part Confidence Maps (PAMs)와 관절과 관절 사이 연결 관계를 표현하는 Part Affinity Fields(PAFs)를 얻는다. 그리고 관절이 어떤 사람의 것인지 판단하기 위해 NP-Hard 매칭 문제를 Greedy 방법으로 풀어 빠르면서도 효과적으로 해결한다[5]. 그림 2는 RGB 입력 영상에 대한 PAMs, PAFs, 추정된 관절의 시각화 이다.

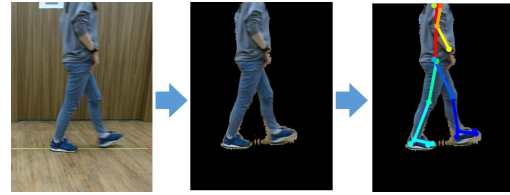


그림 2. Tusimple DUC와 Openpose 알고리즘을 융합하여 보행 자세 골격 구조 추출 결과. (위)원본, (가운데)객체 segmentation 결과, (아래) skeleton 추출 결과

III. 결 론

본 논문에서는 보행 파라미터 획득을 위한 2가지의 보행 자세 검출 접근법을 비교하였다. BMC 을 이용한 방법의 경우, 인접한 픽셀값의 최소값 최대값을 연이어 수행하는 연산방법으로 이웃한 픽셀값에 따라 결과가 달라지며 영상의 사이즈가 클 경우 시간이 오래 걸리는 단점이 있다. 반면 객체 인식 Segmentation 알고리즘과 자세 검출 알고리즘을 융합한 방법을 이용한 접근법의 경우, Bottom-up방식으로 객체의 관절 위치를 탐색하고 관절간의 연결하는 접근방법으로 첫 번째 접근법에 비해 주변 환경 영향이 적다. 두 번째 접근법은 보행 데이터 취득 시, 주변 환경에 상관없이 보행 자세 골격 구조 추출이 가능하였다.

향후 두 번째 접근법은 다양한 보행 형태를 파악하기 위한 파라미터로서 좋은 요소가 될 것이다.

Acknowledgement

본 연구는 미래창조과학부 및 정보통신기술진흥센터의 정보통신-방송 연구개발 사업의 일환으로 수행하였음. [2017-0-00162, 고령 사회에 대응하기 위한 실환경 휴먼케어 로봇 기술 개발]

References

- [1] H. B. Skinner and D. J. Effeny, "Gait analysis in amputees," *American Journal of Physical Medicine*, Vol. 64, No. 2, pp. 82-89, 198.
- [2] M. P. Griffin, S. J. Olney, and I. D. McBride, "Role of symmetry in gait performance of stroke subjects with hemiplegia," *Gait Posture*, Vol. 3, No. 3, pp. 132-142, 199
- [3] Antoine Vacavant, Thierry Chateau, Alexis Wilhelm, and Laurent Lequière. A benchmark dataset for outdoor foreground/background extraction. In *Computer Vision-ACCV 2012 Workshops*, pages 291 - 300. Springer, 2013.
- [4] Wang, Panqu and Chen, Pengfei and Yuan, Ye and Liu, Ding and Huang, Zehua and Hou, Xiaodi and Cottrell, Garrison, "Understanding convolution for semantic segmentation," arXiv

preprint arXiv:1702.08502, 2017.

- [5] Cao Z. Simon T. Wei S. Sheikh Y. (2017) Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields, 2017 IEEE Conference on Computer Vision and Pattern Recognition.